Ann. Inst. Statist. Math.
28 (1976), Part A, 349-357

ON ASYMPTOTIC PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATES OF THE GENERAL GROWTH CURVE MODEL*

S. R. CHAKRAVORTI

(Received Oct. 29, 1973; revised Mar. 25, 1974)

1. Introduction

The problem of estimation of the parameters of the growth curve model under Behrens-Fisher situation has been discussed in Chakravorti [3]. Here we have considered the most general situation by violating the assumption of normality of the underlying parent distribution.

Let us consider the observation vector $\boldsymbol{Z}_\alpha^{(t)}(1\times q):(\boldsymbol{Y}_\alpha^{(t)}, \boldsymbol{X}_\alpha^{(t)})$, the αth observation in tth population ($\alpha=1,\cdots,n_t$, $t=1,\cdots,m$), where $\boldsymbol{Y}_\alpha^{(t)}$ is $1\times p$ vector and $\boldsymbol{X}_\alpha^{(t)}$ is $1\times s$, $s=q-p\geqq 0$. Then considering growth curve model as MANOCOVA model (Rao [10]) we have

$$\boldsymbol{Y}_\alpha^{(t)}=\boldsymbol{\eta}^{(t)}+\boldsymbol{X}_\alpha^{(t)}\boldsymbol{\beta}+\boldsymbol{\varepsilon}_\alpha^{(t)} \tag{1.1}$$

where $\boldsymbol{\eta}^{(t)}(1\times p)$, $\boldsymbol{\beta}(s\times p)$ are the parameters involved in the model, $\boldsymbol{\varepsilon}_\alpha^{(t)}$, the random error component, distributed with continuous distribution function, $F_p(\boldsymbol{\varepsilon}_\alpha^{(t)})$, $\boldsymbol{X}_\alpha^{(t)}(1\times s)$, the concomitant vector variable distributed as $F_s(\boldsymbol{x}_\alpha^{(t)})$. Our object, here, is to study the asymptotic properties of the maximum likelihood estimates (m.l.e.) of the parameters, if they exist. The asymptotic efficiencies of the estimates have been compared with those of the m.l.e.'s obtained under the normality assumptions.

Inagaki [6] discussed these properties for independent not necessarily identically distributed (i.n.i.d.) random variables with parameters $\boldsymbol{\theta}(k\times 1)$ in the line of Huber [4]. In the model (1.1), we consider observation vectors $\boldsymbol{Y}_\alpha^{(t)}$ as i.i.d. with respect to $\alpha=1,\cdots,n_t$ and i.n.i.d. with respect to $t=1,2,\cdots,m$. Then in line of Inagaki we have established the consistency and asymptotic normality of the estimates of the parameters of the model (1.1) under less restrictive assumptions (that is, without the assumption (vii) of Inagaki [6]).

* Work sponsored by the Aerospace Research Laboratories, U.S. Air Force Systems Command, Contract F336152-71-C-1927. Reproduction in whole or in part permitted for any purpose of the U.S. Government.

2. Notations and assumptions

Let us define $\boldsymbol{\theta}$, the parameter matrix of order $(m+s)\times p$, where

$$\boldsymbol{\theta}'=(\boldsymbol{\eta}^{(1)\prime},\cdots,\boldsymbol{\eta}^{(m)\prime},\boldsymbol{\beta}')\ . \tag{2.1}$$

The likelihood function is given by

$$L=\prod_{t=1}^{m}\prod_{\alpha=1}^{n_t} f(\boldsymbol{Y}_\alpha^{(t)}-\boldsymbol{\eta}^{(t)}-\boldsymbol{X}_\alpha^{(t)}\boldsymbol{\beta})\ . \tag{2.2}$$

GENERAL NOTATION

$(\mathcal{X},\mathcal{A},\mathcal{P})$: probability space,

$\boldsymbol{\Theta}$: a parameter space which is a subset of the k $(=(m+s)p)$ dimensional Euclidean space R^k such that for any $M>0$, $\boldsymbol{\Theta}\cap\{\|\boldsymbol{\theta}\|\leqq M\}$ is closed,

$\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}^{(t)})=\partial\log f(\boldsymbol{Y}_\alpha^{(t)}-\boldsymbol{\eta}^{(t)}-\boldsymbol{X}_\alpha^{(t)}\boldsymbol{\beta})/\partial\boldsymbol{\eta}^{(t)}$, $\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta})=\partial\log f(\boldsymbol{Y}_\alpha^{(t)}-\boldsymbol{\eta}^{(t)}-\boldsymbol{X}_\alpha^{(t)}\boldsymbol{\beta})/\partial\boldsymbol{\beta}$, $(t=1,\cdots,m,\ \alpha=1,\cdots,n_t)$ are functions in $\mathcal{X}\times\boldsymbol{\Theta}$

$\|\cdot\|$: the maximum norm of the matrix,

$\hat{\boldsymbol{\theta}}_n$: maximum likelihood estimate of $\boldsymbol{\theta}$ based on n observations,

$\boldsymbol{T}_n$: any other estimator for $\boldsymbol{\theta}$,

$\mathcal{L}(Y)$, $\mathrm{E}(Y)$, $\mathrm{Cov}(Y)$: the distribution, mean and variance-covariance matrix under probability measure P respectively,

$\mathcal{L}(Y,P)$: distribution of Y under probability measure P which is specified.

We shall make the following assumptions, similar to those of Inagaki [6].

ASSUMPTIONS

(i) $\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}^{(t)})$ and $\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta})$ are $\mathcal{B}$-measurable, where $\mathcal{B}$ is the σ-field of Borel-subsets of $\boldsymbol{\Theta}$ and separable when considered as a process of $\boldsymbol{\theta}$.

(ii)

$$\mathrm{E}\begin{bmatrix}\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}_0^{(t)})\\ \boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta}_0)\end{bmatrix}=\boldsymbol{\lambda}^{(t)}(\boldsymbol{\theta}_0)=\begin{bmatrix}\boldsymbol{\lambda}^{(t)}(\boldsymbol{\eta}_0^{(t)})\\ \boldsymbol{\lambda}^{(t)}(\boldsymbol{\beta}_0)\end{bmatrix}=\mathbf{0} \tag{2.3}$$

for any fixed $\boldsymbol{\theta}_0\in\boldsymbol{\Theta}$.

$$\bar{\boldsymbol{\lambda}}_n=\begin{bmatrix}\boldsymbol{\lambda}(\boldsymbol{\eta}^{(1)})\\ \vdots\\ \boldsymbol{\lambda}(\boldsymbol{\eta}^{(m)})\\ \sum_{t=1}^{m} r_t\boldsymbol{\lambda}^{(t)}(\boldsymbol{\beta})\end{bmatrix}=\boldsymbol{\lambda}(\boldsymbol{\theta})\neq 0 \qquad \text{for } \boldsymbol{\theta}\neq\boldsymbol{\theta}_0\,, \tag{2.4}$$

where n and n_t are large so that $r_t=n_t/n$ finite and bounded away from 0 and 1.

(iii) There are two positive constants λ_∞ and $H_\infty>0$ and positive functions $b^{(t)}(\theta)>0$, $t=1,\cdots,m$ such that

$$\mathrm{E}\,[\operatorname*{Sup}_{\boldsymbol{\theta}}\{\boldsymbol{\phi}_\alpha^{(t)}/b^{(t)}(\theta)\}]<\infty \tag{2.5}$$

$$\varlimsup_{n\to\infty}\ \varlimsup_{\|\boldsymbol{\theta}\|\to\infty}\{\operatorname*{Max}_t b^{(t)}(\theta)/\bar{\lambda}_n\}\leqq 1 \tag{2.6}$$

$$\varliminf_{n\to\infty}\ \varliminf_{\|\boldsymbol{\theta}\|\to\infty}\|\bar{\boldsymbol{\lambda}}_n\|>\lambda_\infty>0 \tag{2.7}$$

$$\mathrm{E}\,\{\varlimsup_{\|\boldsymbol{\theta}\|\to\infty}\{\|\boldsymbol{\phi}_\alpha^{(t)}-\boldsymbol{\lambda}^{(t)}\|\}<1 \tag{2.8}$$

$$\mathrm{E}\,\{\varlimsup_{\|\boldsymbol{\theta}\|\to\infty}[\|\boldsymbol{\phi}_\alpha^{(t)}-\boldsymbol{\lambda}^{(t)}\|/b^{(t)}(\theta)]^2\}<H_\infty\,, \tag{2.9}$$

these last two convergences being uniform for $t=1,\cdots,m$, $\alpha=1,\cdots,n_t$.

(iv) For all n_t and n, $\mathrm{E}\,\|\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}^{(t)})-\boldsymbol{\lambda}^{(t)}(\boldsymbol{\eta}^{(t)})\|^2$, $\mathrm{E}\,\|\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta})-\boldsymbol{\lambda}^{(t)}(\boldsymbol{\beta})\|^2$ exist and

$$\frac{1}{n_t^2}\sum_{\alpha=1}^{n_t}\mathrm{E}\,\|\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}^{(t)})-\boldsymbol{\lambda}^{(t)}(\boldsymbol{\eta}^{(t)})\|^2\to 0 \qquad \text{as } n_t\to\infty$$

$$\frac{1}{n^2}\sum_t\sum_\alpha\mathrm{E}\,\|\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta})-\boldsymbol{\lambda}^{(t)}(\boldsymbol{\beta})\|^2\to 0 \qquad \text{as } n\to\infty\,.$$

(v) Let $u_\alpha^{(t)}=\operatorname*{Sup}_{\|\boldsymbol{\tau}-\boldsymbol{\theta}\|<d}\|\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\tau})-\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\theta})\|$, $t=1,\cdots,m$. Then for some constants H_1, H_2 and $d_0>0$, and for $d\leqq d_0$

$$\mathrm{E}\,(u_\alpha^{(t)})<H_1 d\,, \qquad \mathrm{V}\,(u_\alpha^{(t)})<H_2 d$$

(vi) $$\boldsymbol{\Gamma}_{n_t}(\boldsymbol{\eta}^{(t)})=-\mathrm{E}\left[\sum_{\alpha=1}^{n_t}\frac{\partial\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}^{(t)})}{\partial\boldsymbol{\eta}^{(t)\prime}}\right]; \qquad \bar{\boldsymbol{\Gamma}}(\boldsymbol{\eta}^{(t)})=\frac{\boldsymbol{\Gamma}_{n_t}}{n_t}\to\boldsymbol{\Gamma}(\boldsymbol{\eta}^{(t)})$$

$$\boldsymbol{\Gamma}_n(\boldsymbol{\beta})=-\mathrm{E}\left[\sum\sum\frac{\partial\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]; \qquad \bar{\boldsymbol{\Gamma}}(\boldsymbol{\beta})=\frac{\boldsymbol{\Gamma}_n(\boldsymbol{\beta})}{n}\to\boldsymbol{\Gamma}(\boldsymbol{\beta})$$

$$\boldsymbol{\Gamma}_{\eta^{(t)},\beta}=-\mathrm{E}\left[\sum_\alpha\frac{\partial\boldsymbol{\phi}_\alpha^{(t)\prime}(\boldsymbol{\eta}^{(t)})}{\partial\boldsymbol{\beta}}\right]; \qquad \bar{\boldsymbol{\Gamma}}_{\eta^{(t)},\beta}\to\boldsymbol{\Gamma}(\boldsymbol{\eta}^{(t)},\boldsymbol{\beta})\,,$$

uniformly in the neighbourhood of $\boldsymbol{\theta}_0$. Let $\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)$ be the information matrix corresponding to the parameters $\boldsymbol{\eta}_0^{(1)},\cdots,\boldsymbol{\eta}_0^{(m)}$, $\boldsymbol{\beta}_0$ and it is finite and non-singular. It may be noted that derivatives of $\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}^{(t)})$ w.r.t. $\boldsymbol{\eta}^{(t')}$ for $t\neq t'$ are null matrices so that submatrices in the information matrix corresponding to $\boldsymbol{\eta}^{(t)}$, $\boldsymbol{\eta}^{(t')}$ are all null matrices.

(vii) Since the derivatives of the log likelihood function w.r.t. $\boldsymbol{\eta}^{(t)}$ and $\boldsymbol{\beta}$ involve the random vector $\boldsymbol{X}_\alpha^{(t)}$ implicitly, in general, we consider all the expectations as unconditional taking into consideration the following facts:

$$\mathrm{E}(X_\alpha^{(t)})=0\,, \qquad \alpha=1,\cdots,n_t,\ t=1,\cdots,m$$

$$\mathrm{E}(X_\alpha^{(t)\prime}X_\alpha^{(t)})=\Sigma_t^*\,, \qquad \text{finite and positive definite}\,.$$

3. Consistency and asymptotic normality of the maximum likelihood estimates

Under the set up in Section 2 we consider the estimating function $\xi_n(\theta)$ as follows. Let

$$\Delta_{n_t}(\eta^{(t)})=\frac{1}{\sqrt{n_t}}\sum_{\alpha=1}^{n_t}\phi_\alpha^{(t)}(\eta^{(t)})\,, \qquad t=1,\cdots,m \tag{3.1}$$

$$\Delta_n(\beta)=\frac{1}{\sqrt{n}}\sum_{t=1}^{m}\sum_{\alpha=1}^{n_t}\phi_\alpha^{(t)}(\beta)\,. \tag{3.2}$$

Then the estimating function $\xi_n(\theta)$ is given by

$$\xi_n'(\theta)=(\Delta_{n_1}',\cdots,\Delta_{n_m}',\Delta_n')\,. \tag{3.3}$$

Let $\hat{\theta}_n'=(\hat{\eta}^{(1)\prime},\cdots,\hat{\eta}^{(m)\prime},\beta')$ be the maximum likelihood estimates of $\theta'=(\eta^{(1)\prime},\cdots,\eta^{(m)\prime},\beta')$. Then we must have

$$\mathrm{P}\{\lim_{n\to\infty}[\xi_n(\hat{\theta}_n)]=0\}=1\,. \tag{3.4}$$

LEMMA 3.1. *Under the assumptions* (i), (ii), (iv), (vi) *and* (vii),

(a) $$\begin{bmatrix} \frac{1}{\sqrt{n_1}}\Delta_{n_1}(\eta_0^{(1)}) \\ \vdots \\ \frac{1}{\sqrt{n_m}}\Delta_{n_m}(\eta_0^{(m)}) \\ \frac{1}{\sqrt{n}}\Delta_n(\beta_0) \end{bmatrix} \to 0 \ \textit{in}\ \mathrm{P}$$

(b) $[\xi_n(\theta_0)]\to N(0,\Gamma(\theta_0))$ *in law.*

PROOF. (a) From the assumptions (i), (ii), (iv) and (vii) each of $(1/\sqrt{n_t})\Delta_{n_t}(\eta_0^{(t)})$, $t=1,\cdots,m$, converges to zero in probability and since m is finite, r_t bounded, $(1/\sqrt{n_t})\sum_{\alpha=1}^{n_t}\phi_\alpha^{(t)}(\beta_0)$ converges to zero, it implies that $(1/\sqrt{n})\Delta_n(\beta_0)$ converges to zero in probability by W.L.L. NS. (Loeve [9], p. 274). Hence the result follows immediately.

(b) We consider a matrix $\mathcal{H}\in R^k$, where $k=(m+s)p$, such that

$$\mathcal{H}'=[H_1',\cdots,H_m',H_{m+1}']$$

where H_t is of order $1\times p$, $t=1,\cdots,m$ and H_{m+1} of order $s\times p$. Then let us consider the linear function

$$
\begin{aligned}
T &= \operatorname{tr} \boldsymbol{\xi}_n'(\boldsymbol{\theta}_0)\mathcal{H} \\
&= \operatorname{tr} \sum_{t=1}^{m} \boldsymbol{\Delta}_{n_t}' \boldsymbol{H}_t + \operatorname{tr} \boldsymbol{\Delta}_n' \boldsymbol{H}_{m+1} \\
&= \operatorname{tr} \sum_{t=1}^{m} \frac{1}{\sqrt{n_t}} \sum_{\alpha=1}^{n_t} \boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}_0^{(t)})' \boldsymbol{H}_t + \operatorname{tr} \frac{1}{\sqrt{n}} \sum_t \sum_\alpha \boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta}_0)' \boldsymbol{H}_{m+1} \\
&= \sum_t \frac{1}{\sqrt{n_t}} \operatorname{tr} \sum_{\alpha=1}^{n_t} \{\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}_0^{(t)})' H_t + \sqrt{r_t}\, \boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta}_0)' H_{m+1}\} \\
&= \sum_t \frac{1}{\sqrt{n_t}} \sum_{\alpha=1}^{n_t} U_\alpha^{(t)}
\end{aligned}
$$

where

$$
U_\alpha^{(t)} = \operatorname{tr} (\boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\eta}_0^{(t)})', \sqrt{r_t}\, \boldsymbol{\phi}_\alpha^{(t)}(\boldsymbol{\beta}_0)') \begin{pmatrix} H_t \\ H_{m+1} \end{pmatrix} .
$$

Now from assumptions (ii) and (vi) we have

$$
\mathrm{E}\, U_\alpha^{(t)} = 0
$$

$$
\operatorname{Var}(U_\alpha^{(t)}) = \operatorname{tr} \boldsymbol{\Gamma}^{(t)} \boldsymbol{H}^{(t)} \boldsymbol{H}^{(t)\prime} = G^{(t)} \qquad \text{(say)}
$$

where $\boldsymbol{H}^{(t)\prime} = (\boldsymbol{H}_t', \boldsymbol{H}_{m+1}')$ and

$$
\boldsymbol{\Gamma}^{(t)} = \begin{bmatrix} \boldsymbol{\Gamma}(\boldsymbol{\eta}_0^{(t)}) & \sqrt{r_t}\, \boldsymbol{\Gamma}(\boldsymbol{\eta}_0^{(t)}, \boldsymbol{\beta}_0) \\ \sqrt{r_t}\, \boldsymbol{\Gamma}'(\boldsymbol{\eta}_0^{(t)}, \boldsymbol{\beta}_0) & r_t \boldsymbol{\Gamma}^{(t)}(\boldsymbol{\beta}_0) \end{bmatrix}
$$

is finite. Hence $\{U_\alpha^{(t)}\}$ satisfies the Lindeberg-Levy's condition for C.L.T., so that

$$
\mathcal{L}\left[\frac{1}{n_t} \sum_{\alpha=1}^{n_t} U_\alpha^{(t)}\right] \to N(0, G^{(t)}) \text{ in law} .
$$

Hence it follows that $T = \sum_{t=1}^{m} \left(\frac{1}{\sqrt{n_t}} \sum_{\alpha=1}^{n_t} U_\alpha^{(t)}\right)$ is asymptotically distributed as $N(0, G)$, where G is the variance of T, given by $\operatorname{tr} \mathcal{H}' \boldsymbol{\Gamma}(\boldsymbol{\theta}_0) \mathcal{H}$, where $\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)$ is the variance-covariance matrix of $\boldsymbol{\xi}_n(\boldsymbol{\theta}_0)$.

Since for any $\mathcal{H} \in R^k$ this result holds we have the result (b). Hence the lemma is completely proved.

Consistency of the m.l.e. $\hat{\boldsymbol{\theta}}_n$

$\hat{\boldsymbol{\theta}}_n$ being the maximum likelihood estimate for $\boldsymbol{\theta}$, it follows from (3.4) that

$$\begin{bmatrix} \frac{1}{\sqrt{n_1}} \Delta_{n_1}(\hat{\eta}^{(1)}) \\ \vdots \\ \frac{1}{\sqrt{n_m}} \Delta_{n_m}(\hat{\eta}^{(m)}) \\ \frac{1}{\sqrt{n}} \Delta_n(\hat{\beta}) \end{bmatrix} \to 0 \text{ in P} .$$

Hence in the line of Huber [4] and under the assumptions (i)–(v) and (vii), $\{\hat{\theta}_n\}$ converges to θ_0 in P.

Asymptotic normality

To obtain the asymptotic distribution of the m.l.e.'s $\hat{\theta}_n$ in the line of Inagaki [6] let us define the concept of relative compactness (see LeCam [8]) in the following sense.

DEFINITION. In order that $\{\mathcal{L}(Y_n)\}$ is said to be relatively compact it is necessary and sufficient that for any $\varepsilon>0$ there exists a positive number $M>0$ such that

$$\text{(3.5)} \qquad P\{\|Y_n\|>M\}<\varepsilon , \qquad \text{for all } n .$$

Now $\hat{\theta}_n$ being m.l.e. of θ_0, the relation (3.4) implies that $\{\mathcal{L}(\xi_n(\hat{\theta}_n))\}$ is relatively compact. Hence from Theorem 3.2 of Inagaki [6], $\{\mathcal{L}[\sqrt{n}\cdot(\hat{\theta}_n-\theta_0)]\}$ is relatively compact. Hence we have the following.

THEOREM 3.1. *Under the assumptions* (i)–(iii) *if* $\{\mathcal{L}[\sqrt{n}(\hat{\theta}_n-\theta_0)]\}$ *is relatively compact then*

$$\text{(3.6)} \qquad \delta_n(\sqrt{n}(\hat{\theta}_n-\theta_0))=\xi_n(\hat{\theta}_n)-\xi_n(\theta_0)+\sqrt{n}(\hat{\theta}_n-\theta_0)\Gamma\to 0 \text{ in P}$$

where

$$\text{(3.7)} \qquad \sqrt{n}(\hat{\theta}_n-\theta_0)=\begin{bmatrix} \sqrt{n_1}(\hat{\eta}^{(1)}-\eta_0^{(1)}) \\ \vdots \\ \sqrt{n_m}(\hat{\eta}^{(m)}-\eta_0^{(m)}) \\ \sqrt{n}(\hat{\beta}-\beta_0) \end{bmatrix} .$$

PROOF. Following Lemma 3.2 of Inagaki [6], page 7, it can be easily shown that, under the assumptions (i), (ii), (v) and (vi), for any $M>0$ and large n (putting $T=\sqrt{n}(\tau-\theta_0)$),

$$\text{(3.8)} \qquad \sup_{\|T\|\le M}\|\delta_n(T)\|=\sup_{\|T\|\le M}\left\|\xi_n\left(\theta_0+\frac{T}{\sqrt{n}}\right)-\xi_n(\theta_0)-T\Gamma(\theta_0)\right\|\to 0 \text{ in P} .$$

This shows that $\xi_n(\theta)$ is "weakly asymptotically differentiable" (in the sense of Inagaki [6]). Now since $\{\mathcal{L}[\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)]\}$ is relatively compact, the result (3.6) follows from (3.8) in the line of Theorem 3.1 of of Inagaki [6]. Hence the theorem.

Thus we have

THEOREM 3.2. *Under the assumptions* (i)-(vii)

$$\mathcal{L}[\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)]\rightarrow N(0, \boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)) \text{ in law .}$$

PROOF. Since (3.4) holds for m.l.e.'s $\hat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}_0$, we have from Theorem 3.1 that

$$\mathcal{L}[\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)]\rightarrow \mathcal{L}[\boldsymbol{\xi}_n(\boldsymbol{\theta}_0)] .$$

Hence from Lemma 3.1 (b) we have the desired result and the theorem is completely proved.

4. Asymptotic efficiency of $\hat{\boldsymbol{\theta}}_n$

$\hat{\boldsymbol{\theta}}_n$ being m.l.e.'s of the parameter matrix $\boldsymbol{\theta}_0$, is complete sufficient for $P_{\boldsymbol{\theta}_0+\mathcal{H}/\sqrt{n}}$, where

$$\frac{\mathcal{H}}{\sqrt{n}}=\left(\frac{1}{\sqrt{n_1}}\boldsymbol{H}_1',\cdots,\frac{1}{\sqrt{n_m}}\boldsymbol{H}_m',\frac{1}{\sqrt{n}}\boldsymbol{H}_{m+1}'\right) \tag{4.1}$$

so that

$$\boldsymbol{\theta}_0+\frac{\mathcal{H}}{\sqrt{n}}=\begin{bmatrix} \boldsymbol{\eta}_0^{(1)}+\dfrac{\boldsymbol{H}_1}{\sqrt{n_1}} \\ \vdots \\ \boldsymbol{\eta}_0^{(m)}+\dfrac{\boldsymbol{H}_m}{\sqrt{n_m}} \\ \boldsymbol{\beta}_0+\dfrac{\boldsymbol{H}_{m+1}}{\sqrt{n}} \end{bmatrix} . \tag{4.2}$$

Then any other estimator $\boldsymbol{T}_n$ which is location invariant is independent of $\hat{\boldsymbol{\theta}}_n$ (see Basu [1], [2]).

Now $\boldsymbol{T}_n$ will be said to be asymptotically location invariant at $\boldsymbol{\theta}_0$ if for any $\mathcal{H}\in R^k$

$$\mathcal{L}\left[\boldsymbol{T}_n-\left(\boldsymbol{\theta}_0+\frac{\mathcal{H}}{\sqrt{n}}\right);\, P_{\boldsymbol{\theta}_0+\mathcal{H}/\sqrt{n}}\right]\rightarrow L \text{ in law} \tag{4.3}$$

where L is independent of $\mathcal{H}$. The necessary and sufficient condition for this is that $\boldsymbol{\xi}_n(\boldsymbol{T}_n)$ is location invariant at $\boldsymbol{\theta}_0$. That is

$$\mathcal{L}[\boldsymbol{\xi}_n(\boldsymbol{T}_n);\, P_{\boldsymbol{\theta}_0+\mathcal{H}/\sqrt{n}}]\rightarrow G \text{ in law} \tag{4.4}$$

where G is independent of $\mathcal{H}$.

Under this set up, following Theorem 5.1 of Inagaki [6], it can be shown that the limiting distribution L can be expressed as a convolution of those of $\boldsymbol{\xi}_n(\boldsymbol{\theta}_0)$ and $-\boldsymbol{\xi}_n(\boldsymbol{T}_n)$ under $P_{\boldsymbol{\theta}_0}$, $N(0, \boldsymbol{\Gamma}(\boldsymbol{\theta}_0))$ and $\tilde{G}(Z)=1-G(-Z)$, that is

$$L=\tilde{G} * N(0, \boldsymbol{\Gamma}) . \tag{4.5}$$

Hence it follows from Corollary 5.1 of Inagaki [6] that

$$\lim [\sqrt{n}(\boldsymbol{T}_n-\boldsymbol{\theta}_0)]=\lim [\sqrt{n}(\boldsymbol{T}_n-\hat{\boldsymbol{\theta}}_n)] * \lim [\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)] . \tag{4.6}$$

Since $\hat{\boldsymbol{\theta}}_n$ is sufficient statistic for $\boldsymbol{\theta}_0$ it follows from Kaufman [7] that $(\boldsymbol{T}_n-\hat{\boldsymbol{\theta}}_n)$ and $\hat{\boldsymbol{\theta}}_n$ are independently distributed. Hence if $\boldsymbol{V}^{(1)}$ and $\boldsymbol{V}^{(2)}$ be the dispersion matrices corresponding to $\boldsymbol{T}_n$ and $\hat{\boldsymbol{\theta}}_n$, then it follows from (4.6) that $\boldsymbol{V}^{(1)}-\boldsymbol{V}^{(2)}$ must be at least positive semidefinite. This proves that $\hat{\boldsymbol{\theta}}_n$ is asymptotically efficient as compared with the estimator which is asymptotically l-invariant.

5. Relative efficiency of $\hat{\boldsymbol{\theta}}_n$ compared with m.l.e. when parent distribution is multinormal

The m.l. estimators of the parameters of the growth curve model under Behrens-Fisher situation have been considered in [3], so that when the underlying distribution is multinormal the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)$, defined by (3.7), has been shown to be $N(0, \mathcal{I}^{-1})$ where $\mathcal{I}$ is the information matrix with off-diagonal submatrices zero.

To compare the relative efficiency of the estimates $\hat{\boldsymbol{\theta}}_n$ with that under normality we are only to compare the corresponding dispersion matrices, $\boldsymbol{V}^{(1)}(\boldsymbol{\theta}_0)=\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)$ and $\boldsymbol{V}^{(2)}(\boldsymbol{\theta}_0)=\mathcal{I}^{-1}(\boldsymbol{\theta}_0)$, so that relative efficiency is given by

$$e=\left\{\frac{|\boldsymbol{V}^{(2)}|}{|\boldsymbol{V}^{(1)}|}\right\}^{1/p} . \tag{5.1}$$

This $e\leqq 1$ provided $\boldsymbol{V}^{(1)}-\boldsymbol{V}^{(2)}$ is at least positive semidefinite.

Acknowledgement

The author is grateful to Professor P. K. Sen for his help and guidance throughout this investigation.

UNIVERSITY OF NORTH CAROLINA*

* On leave of absence from the University of Calcutta.

REFERENCES

[1] Basu, D. (1955). On statistics independent of a complete sufficient statistic, *Sankhya*, **15**, 377-380.

[2] Basu, D. (1958). On statistics independent of sufficient statistics, *Sankhya*, **18**, 223-226.

[3] Chakravorti, S. R. (1975). Maximum likelihood estimation of growth curve model with unequal dispersion matrices, *Commun. Statist.*, **4**, 69-85.

[4] Huber, P. J. (1967)· The behaviour of maximum likelihood estimators under non-standard conditions, *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, **1**, 221-233.

[5] Inagaki, N. (1970). On the limiting distribution of a sequence of estimators with uniformity property, *Ann. Inst. Statist. Math.*, **22**, 1-13.

[6] Inagaki, N. (1973). Asymptotic relation between the likelihood estimating function and the maximum likelihood estimator, *Ann. Inst. Statist. Math.*, **25**, 1-25.

[7] Kaufman, S. (1966). Asymptotic efficiency of the maximum likelihood estimator, *Ann. Inst. Statist. Math.*, **18**, 155-178.

[8] LeCam, L. (1960). Locally asymptotically normal families of distributions, *Univ. Calif. Pub. Statist.*, **3**, 37-98.

[9] Loeve, L. (1963). *Probability Theory*, Von Nostrand, Princeton.

[10] Rao, C. R. (1965). The theory of least squares when parameters are stochastic and its applications to the analysis of growth curve, *Biometrika*, **52**, 447-458.